

SEQUENCE LISTING

SEQ ID NO:1

Mouse SSG amino acid sequence

5 MGELPFLSPEGARGPHINRGSLSSLEQGSVTGTEARHSLGVLHVSYSVSNRVGPWWNIKS
CQQKWDRQILKDVSPLYIESGQIMCILGSSSGSGKTTLLDAISGRLRRTGTLEGEV FVNGCE
LRDQFQDCFSYVLQSDVFLSSLTVRETLRYTAMLALCRSSADFYNNKKVEAVMTELSLSH
VADQMIGSYNFGGISSGERRRVSIAAQLLQDPKVMMLDEPTTGLDCMTANQIVLLLAEAL
RRDRIVIVTIIHQPRSELFQHFDKIAILTYGELVFCGTPEEMLGFFNNCGYPCPEHSNPFD
10 FYMDLTSVDTQSREREIETYKRVQMLECAFKESDIYHKILENIERARYLKTLPMPFKTK
DPPGMFGKLGVLRRVTRNLMRNKQAVIMRLVQNLIMGLFLIFYLLRVQNNTLKGAVQDR
VGLLYQLVGATPYTGMLNAVNLFPMLRAVSDQESQDGLYHKWQMLLAYVLHVLPPFSVIAT
VIFSSVCYWTGLGLYPEVARFGYFSAALLAPHLIGEFLLVLLGIVQNPNI VNSIVALLSI
SGLLIGSGFIRNIQEMPIPLKILGYFTFQKYCCEILVVNEFYGLNFTCGGSNTSMLNHPM
15 CAITQGVQFIEKTCPGATSRFTANFLILYGFIPALVILGIVIFKVRDYLI SR

SEQ ID NO:2

Mouse SSG nucleotide sequence

20 GGGACAGGCCACTAGAAAATTCACCTGCATTTGCTTCCTGCTAGCCATGGGTGAGCTGCC
CTTTCTGAGTCCAGAGGGAGCCAGAGGGCCTCACATCAACAGAGGGTCTCTGAGCTCCCT
GGAGCAAGGTTTCGGTCACGGGCACAGAGGCTCGGCACAGCTTAGGTGTCCTGCATGTGTC
CTACAGCGTCAGCAACCGTGTCTGGGCCTTGGTGGAACATCAAATCATGCCAGCAGAAGTG
GGACAGGCAAATCCTCAAAGATGTCTCCTTGACATCGAGAGTGGCCAGATTATGTGCAT
25 CTTAGGCAGCTCAGGCTCAGGGAAGACCACGCTGCTGGACGCCATCTCCGGGAGGCTGCG
GCGCACTGGGACCCTGGAAGGGGAGGTGTTTGTGAATGGCTGCGAGCTGCGCAGGGACCA
GTTCCAAGACTGCTTCTCCTACGTCCTGCAGAGCGACGTTTTTCTGAGCAGCCTCACTGT
GCGCGAGACGTTGCGATACACAGCGATGCTGGCCCTCTGCCGCAGCTCCGCGGACTTCTA
CAACAAGAAGGTAGAGGCAGTCATGACAGAGCTGAGCCTGAGCCACGTGGCGGACCAAAT
30 GATTGGCAGCTATAATTTTGGGGGAATTTCCAGTGGCGAGCGGCGCCGAGTTTCCATCGC
AGCCCAACTCCTTCAGGACCCCAAGGTCATGATGCTAGATGAGCCAACCACAGGACTGGA
CTGCATGACTGCAAATCAAATTGTCTTCTCTTGGCTGAGCTGGCTCGCAGGGACCGAAT
TGTGATTGTCACCATCCACCAGCCTCGCTCTGAGCTCTTCCAACACTTCGACAAAATTGC
CATCCTGACTTACGGAGAGTTGGTGTCTGTGGCACCCAGAGGAGATGCTTGGCTTCTT

CAATAACTGTGGTTACCCCTGTCTGAACATTCCAATCCCTTTGATTTTTACATGGACTT
GACATCAGTGGACACCCAAAGCAGAGAGCGGGAAATAGAAACGTACAAGCGAGTACAGAT
GCTGGAATGTGCCTTCAAGGAATCTGACATCTATCACAAAATTCTGGAGAACATTGAAAG
AGCACGATACCTGAAAACCTTACCCATGGTTCCTTTCAAAACAAAAGATCCTCCTGGGAT
5 GTTCGGCAAGCTTGGTGTCTGCTGAGGCGAGTAACAAGAACTTAATGAGGAATAAGCA
GGCAGTGATTATGCGTCTCGTTCAGAATCTGATCATGGGCCCTCTTCCTCATTTTTCTACCT
TCTCCGCGTCCAGAACAACACGCTAAAGGGCGCTGTGCAGGACCGCGTGGGGCTGCTCTA
TCAGCTTGTGGGTGCCACCCCATACACCGGCATGCTCAATGCTGTGAATCTGTTTCCCAT
GCTGAGAGCCGTCAGCGACCAGGAGAGTCAGGATGGCCTGTATCATAAGTGGCAGATGCT
10 GCTCGCCTACGTGCTACACGTCCTCCCCTTCAGCGTCATCGCCACGGTCATTTTCAGCAG
TGTGTGTTATTGGACTCTGGGCTTGTATCCTGAAGTTGCCAGATTGGATATTTCTCTGC
TGCTCTTTTGGCCCCCTCACTTAATTGGAGAATTTCTAACACTTGTGCTGCTTGGTATAGT
CCAAAACCCTAATATTGTCAACAGTATAGTGGCTCTGCTCAGCATCTCTGGGCTGCTTAT
TGGATCTGGATTTATCAGAAACATACAAGAAATGCCATTCTTTAAAAATCCTGGGTTA
15 TTTTACATTCCAAAAATACTGTTGTGAGATTCTCGTGGTCAATGAGTTTTACGGCCTGAA
CTTCACTTGTGGTGGATCCAACACCTCTATGCTAAATCACCCGATGTGCGCCATCACCCA
AGGGGTCCAGTTCATCGAGAAAACCTGCCAGGTGCTACATCCAGATTACGGCAAACCTT
CCTCATCTTATATGGGTTTATCCCAGCTCTGGTCATCCTAGGAATAGTGATTTTTAAAGT
CAGGGACTACCTGATTAGCAGATAGTTAAGATGACAGGCAGGAAAGGGTTAATGGGCAGG
20 CACGCCCCACTGTGGAGCACAGAGAAGTACTGTCTTCAACCATCAGGATTCCATCTGCGAC
CCTTGTGTCTGACCCTTGTGTCTATCCGGAGCCCCAAGGGCAACGAGAACTCACAGCCCT
CTGCTATTCCAGCTTGTGGGGCAATGTGGTGCTTGGACATTGTGACTGAACTGGTCCAAT
AATGTAAATAATAATAATTCATAAACCTACAGGACATT

SEQ ID NO:3

Human SSG amino acid sequence

MGDLSSLTPGGSMGLQVNRGSQSSLEGAPATAPEPHSLGILHASYSVSHRVRPWWDTSC
30 RQQWTRQILKDVS LYVESGQIMCILGSSGSGKTTLLDAMSGRLGRAGTFLGEVYVNGRAL
RREQFQDCFSYVLQSDTLLSSLTVRETLHYTALLAIRRGNPGSFQKKVEAVMAELSLSHV
ADRLIGNYSLGGISTGERRRVSIQAQLLQDPKVMLFDEPTTGLDCMTANQIVVLLVELAR
RNRIVVLTIHQPRSELFQLFDKIAILSFGELIFCGTPAEMLDFFNDCGYPCPEHSNPFDF
YMDLTSVDTQSKEREIETSKRVQMIESAYKKS AICHKTLKNIERMKHLKTLPMVPFKTKD

SPGVFSKLGVLRLRRVTRNLVRNKLAVITRLLQNLIMGLFLLFFVLRVRSNVLKGAIQDRV
 GLLYQFVGATPYTGMLNAVNLFPVLRVSDQESQDGLYQKWQMMLAYALHVLFPFSVVATM
 IFSSVCYWTLGLHPEVARFGYFSAALLAPHLIGEFLLVLLGIVQNPNI VNSVVALLSIA
 GVLVSGSGLRNIQEMPIPFKIIISYFTFQKYCSEILVVNEFYGLNFTCGSSNVSVTTNPMC
 5 AFTQGIQFIEKTCPGATSRFTMNFLILYSFIPALVILGIVVFKIRDHLISR

SEQ ID NO:4

Human SSG nucleotide sequence

10

GTCAGGTGGAGCAGGCAGGGCAGTCTGCCACGGGCTCCCCAACTGAAGCCACTCTGGGGA
 GGGTCCGGCCACCAGAAAAATTTGCCAGCTTTGCTGCCTGTTGGCCATGGGTGACCTCTC
 ATCTTTGACCCCCGGAGGGTCCATGGGTCTCCAAGTAAACAGAGGCTCCCAGAGCTCCCT
 GGAGGGGGCTCCTGCCACCGCCCCGGAGCCTCACAGCCTGGGCATCCTCCATGCCTCCTA

15

CAGCGTCAGCCACCGCGTGAGGCCCTGGTGGGACATCACATCTTGCCGGCAGCAGTGGAC
 CAGGCAGATCCTCAAAGATGTCTCCTTGTACGTGGAGAGCGGGCAGATCATGTGCATCCT
 AGGAAGCTCAGGCTCCGGGAAAACCACGCTGCTGGACGCCATGTCCGGGAGGCTGGGGCG
 CGCGGGGACCTTCTGGGGGAGGTGTATGTGAACGGCCGGGCGCTGCGCCGGGAGCAGTT
 CCAGGACTGCTTCTCCTACGTCCTGCAGAGCGACACCCTGCTGAGCAGCCTACCGTGCG

20

CGAGACGCTGCACTACACCGCGCTGCTGGCCATCCGCCGCGGCAATCCCGGCTCCTTCCA
 GAAGAAGGTGGAGGCCGTCATGGCAGAGCTGAGTCTGAGCCATGTGGCAGACCGACTGAT
 TGGCAACTACAGCTTGGGGGGCATTTCACGGGTGAGCGGCGCGGGTCTCCATCGCAGC
 CCAGCTGCTCCAGGATCCTAAGGTCATGCTGTTTGATGAGCCAACCACAGGCCTGGACTG
 CATGACTGCTAATCAGATTGTGCTCCTCCTGGTGGAACTGGCTCGCAGGAACCGAATTGT

25

GGTTCACCATTCACCAGCCCCGTTCTGAGCTTTTTTCAGCTCTTTGACAAAATTGCCAT
 CCTGAGCTTCGGAGAGCTGATTTTCTGTGGCACGCCAGCGGAAATGCTTGATTTCTTCAA
 TGACTGCGGTTACCCTTGTCCTGAACATTCAAACCTTTTGACTTCTATATGGACCTGAC
 GTCAGTGGATACCCAAAGCAAGGAACGGGAAATAGAAACCTCCAAGAGAGTCCAGATGAT
 AGAATCTGCCTACAAGAAATCAGCAATTTGTCATAAACTTTGAAGAATATTGAAAGAAT

30

GAAACACCTGAAAACGTTACCAATGGTTCCTTTCAAACCAAAGATTCTCCTGGAGTTTT
 CTCTAAACTGGGTGTTCTCCTGAGGAGAGTGACAAGAACTTGGTGAGAAATAAGCTGGC
 AGTGATTACGCGTCTCCTTCAGAATCTGATCATGGGTTTGTTCCTCCTTTTCTTCGTTCT
 GCGGGTCCGAAGCAATGTGCTAAAGGGTGCTATCCAGGACCGGTAGGTCTCCTTTACCA
 GTTTGTGGGCGCCACCCCGTACACAGGCATGCTGAACGCTGTGAATCTGTTTCCCGTGCT

5 GCGAGCTGTCAGCGACCAGGAGAGTCAGGACGGCCTCTACCAGAAGTGGCAGATGATGCT
GGCCTATGCACTGCACGTCTCTCCCTTCAGCGTTGTTGCCACCATGATTTTCAGCAGTGT
GTGCTACTGGACGCTGGGCTTACATCCTGAGGTTGCCCGATTTGGATATTTTTCTGCTGC
TCTCTTGGCCCCCCTTAATTGGTGAATTTCTAACTCTTGTGCTACTTGGTATCGTCCA
10 AAATCCAAATATAGTCAACAGTGTAGTGGCTCTGCTGTCCATTGCGGGGGTGCTTGTG
ATCTGGATTCTCAGAAACATACAAGAAATGCCCATTCCTTTTAAAATCATCAGTTATTT
TACATTCCAAAAATATTGCAGTGAGATTCTTGTAGTCAATGAGTTCTACGGACTGAATTT
CACTTGTGGCAGCTCAAATGTTTCTGTGACAATAATCCAATGTGTGCCTTCACTCAAGG
AATTCAATTCATTGAGAAAACCTGCCCAGGTGCAACATCTAGATTCACAATGAACTTTCT
15 GATTTTGTATTCAATTTATTCCAGCTCTTGTATCCTAGGAATAGTTGTTTTCAAAATAAG
GGATCATCTCATTAGCAGGTAGTGAAAGCCATGGCTGGGAAAATGGAAGTGAAGCTGCCG
ACTGTGCATGACTGCTCTGAACGTCTGAAATGAGAGTGCCATGTATTTCTTTCTTGACAG
GACATCTCAAGTCTTTTAACCATTAAGACTCCATTTGTGCCTCTTGGATCCAAGCAGGCC
TTGAATGCAATGGAAGTGGTTTATAGTCCCTTGCTCTTACAACCTGCAGGGACATGTGGT
TATTTGGAAATTGTGACTGAGCGGACCCAAGAATGTAAATAATATTCATAAACCTATGGG

SEQ ID NO:5

SSG signature sequence 1

20

AALLAPHLIGEFLTLVLL

SEQ ID NO:6

25 SSG signature sequence 2

FIPALVILGIV

SEQ ID NO:7

30 Exon 1 of hSSG

GTCAGGTGGAGCAGGCAGGGCAGTCTGCCACGGGCTCCCCAACTGAAGCCACTCTGGGGA
GGGTCCGGCCACCAGAAAATTTGCCCAGCTTTGCTGCCTGTTGGCCATGGGTGACCTCTC
ATCTTTGACCCCCGGAGGGTCCATGGGTCTCCAAGTAAACAGAGGCTCCCAGAGCTCCCT

GGAGGGGGCTCCTGCCACCGCCCCGGAGCCTCACAGCCTGGGCATCCTCCATGCCTCCTA
CAGCGTCAG

5 **SEQ ID NO:8**

Exon 2 of hSSG

CCACCGCGTGAGGCCCTGGTGGGACATCACATCTTGCCGGCAGCAGTGGACCAGGCAGAT
CCTCAAAGATGTCTCCTTGTACGTGGAGAGCGGGCAGATCATGTGCATCCTAGGAAGCTC

10 AG

SEQ ID NO:9

Exon 3 of hSSG

15 GCTCCGGGAAAACCACGCTGCTGGACGCCATGTCCGGGAGGCTGGGGCGCGCGGGGACCT
TCCTGGGGGAGGTGTATGTGAACGGCCGGGCGCTGCGCCGGGAGCAGTTCCAGGACTGCT
TCTCCTACGTCTTGCAG

SEQ ID NO:10

20 Exon 4 of hSSG

AGCGACACCCTGCTGAGCAGCCTCACCGTGCGCGAGACGCTGCACTACACCGCGCTGCTG
GCCATCCGCCGCGGCAATCCCGGCTCCTTCCAGAAGAAGGTGG

25 **SEQ ID NO:11**

Exon 5 of hSSG

AGGCCGTCATGGCAGAGCTGAGTCTGAGCCATGTGGCAGACCGACTGATTGGCAACTACA
GCTTGGGGGGCATTTCACGGGTGAGCGGCGCCGGGTCTCCATCGCAGCCCAGCTGCTCC

30 AGGATCCTA

SEQ ID NO:12

Exon 6 of hSSG

bioRxiv preprint doi: <https://doi.org/10.1101/044404>; this version posted April 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AGGTCATGCTGTTTGATGAGCCAACCACAGGCCTGGACTGCATGACTGCTAATCAGATTG
TCGTCCTCCTGGTGGAACTGGCTCGCAGGAACCGAATTGTGGTTCTCACCATTACCAGC
CCCGTTCTGAGCTTTTTCAG

5 **SEQ ID NO:13**

Exon 7 of hSSG

CTCTTTGACAAAATTGCCATCCTGAGCTTCGGAGAGCTGATTTTCTGTGGCAGCCAGCG
GAAATGCTTGATTTCTTCAATGACTGCGGTTACCCTTGTCTGAACATTCAAACCCCTTTT

10 GACTTCTATA

SEQ ID NO:14

Exon 8 of hSSG

15 TGGACCTGACGTCAGTGGATACCCAAAGCAAGGAACGGGAAATAGAAACCTCCAAGAGAG
TCCAGATGATAGAATCTGCCTACAAGAAATCAGCAATTTGTCATAAACTTTGAAGAATA
TTGAAAGAAATGAAACACCTGAAAACGTTACCAATGGTTCCTTTCAAACCAAAGATTCTC
CTGGAGTTTTCTCTAAACTGGGTGTTCTCCTGAG

20 **SEQ ID NO:15**

Exon 9 of hSSG

GAGAGTGACAAGAACTTGGTGAGAAATAAGCTGGCAGTGATTACGCGTCTCCTTCAGAA
TCTGATCATGGGTTTGTTCCTCCTTTTCTTCGTTCTGCGGGTCCGAAGCAATGTGCTAAA
25 GGGTGCTATCCAGGACCGCTAGGTCTCCTTTACCAGTTTGTGGGCGCCACCCCGTACAC
AGGCATGCTGAACGCTGTGAATCTGT

SEQ ID NO:16

Exon 10 of hSSG

30

TTCCCGTGCTGCGAGCTGTCAGCGACCAGGAGAGTCAGGACGGCCTCTACCAGAAGTGGC
AGATGATGCTGGCCTATGCACTGCACGTCCTCCCCTTCAGCGTTGTTGCCACCATGATTT
TCAGCAGTGTGTGCTACTG

SEQ ID NO:17

Exon 11 of hSSG

5 GACGCTGGGCTTACATCCTGAGGTTGCCCGATTTGGATATTTTTCTGCTGCTCTCTTGGC
CCCCCACTTAATTGGTGAATTTCTAACTCTTGTGCTACTTGGTATCGTCCAAAATCCAAA
TATAGTCAACAGTGTAGTGGCTCTGCTGTCCATTGCGGGGGTGCTTGTGGATCTGGATT
CCTCAG

SEQ ID NO:18

10 Exon 12 of hSSG

AAACATACAAGAAATGCCCATTCCTTTTAAAATCATCAGTTATTTTACATTCCAAAAATA
TTGCAGTGAGATTCTTGTAGTCAATGAGTTCTACGGACTGAATTTCACTTGTG

15 **SEQ ID NO:19**

Exon 13 of hSSG

GCAGCTCAAATGTTTCTGTGACAACTAATCCAATGTGTGCCTTCACTCAAGGAATTCAAT
TCATTGAGAAAACCTGCCCAGGTGCAACATCTAGATTACAATGAACTTTCTGATTTTGT
20 ATTCATTTATTCCAGCTCTTGTATCCTAGGAATAGTTGTTTTCAAAAATAAGGGATCATC
TCATTAGCAGGTAGTGAAAGCCATGGCTGGGAAAAATGGAAGTGAAGCTGCCGACTGTGCA
TGACTGCTCTGAACGTCTGAAATGAGAGTGCCATGTATTTCTTTCTTGACAGGACATCTC
AAGTCTTTTAACCATTAAGACTCCATTTGTGCCTCTTGGATCCAAGCAGGCCTTGAATGC
AATGGAAGTGGTTTATAGTCCCTTGCTCTTACAACCTTGCAGGGACATGTGGTTATTTGGA
25 AATTGTGACTGAGCGGACCCAAGAATGTAAATAATATTCATAAACCTATGGG